



Docket No.: 4441-0109P
(PATENT)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:
Kristofer KJORLING et al.

Application No.: 09/987,475

Confirmation No.: 007069

Filed: November 14, 2001

Art Unit: 2654

For: APPARATUS AND METHOD APPLYING
ADAPTIVE SPECTAL WHITENING IN A
HIGH-FREQUENCY RECONSTRUCTION
CODING SYSTEM

Examiner: V. B. Chawan

CLAIM FOR PRIORITY AND SUBMISSION OF DOCUMENTS

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

Applicant hereby claims priority under 35 U.S.C. 119 based on the following prior foreign application filed in the following foreign country on the date indicated:

<u>Country</u>	<u>Application No.</u>	<u>Date</u>
Sweden	0004163-2	November 14, 2000

Application No.: 09/987,475

Docket No.: 4441-0109P

In support of this claim, a certified copy of the said original foreign application is filed herewith.

Dated: November 9, 2005

Respectfully submitted,

#40,439

By

Terrell C. Birch

Registration No.: 19,382

BIRCH, STEWART, KOLASCH & BIRCH, LLP

8110 Gatehouse Road

Suite 100 East

P.O. Box 747

Falls Church, Virginia 22040-0747

(703) 205-8000

Attorney for Applicant

PRV

PATENT- OCH REGISTRERINGSVERKET
Patentavdelningen

Intyg Certificate

Härmed intygas att bifogade kopior överensstämmer med de handlingar som ursprungligen ingivits till Patent- och registreringsverket i nedannämnda ansökan.

This is to certify that the annexed is a true copy of the documents as originally filed with the Patent- and Registration Office in connection with the following patent application.

(71) Sökande Coding Technologies Sweden AB, Stockholm SE
Applicant (s)

(21) Patentansökningsnummer 0004163-2
Patent application number

(86) Ingivningsdatum 2000-11-14
Date of filing

Stockholm, 2005-10-19

För Patent- och registreringsverket
For the Patent- and Registration Office



Hjärdís Segerlund

Avgift
Fee 170:-

ENHANCING PERCEPTUAL PERFORMANCE OF HIGH FREQUENCY RECONSTRUCTION CODING METHODS BY ADAPTIVE FILTERING

5 TECHNICAL FIELD

The present invention relates to source coding systems utilising high frequency reconstruction (HFR) such as Spectral Band Replication, SBR [WO 98/57436] or related methods. It improves performance of both high quality methods (SBR), as well as low quality copy-up methods [U.S. Pat. 5,127,054]. It is applicable to both speech coding and natural audio coding systems.

10 BACKGROUND OF THE INVENTION

In high frequency reconstruction, where a low band is used to extrapolate a high band it is important to have means to control the harmonics of the replicated high band. This is necessary since the harmonics usually are stronger in the low band compared to the high band. An extreme example is a very pronounced harmonic series in the lowband and more or less pure noise in the high band. One way to approach this is by adding noise adaptively to the highband (Adaptive Noise Addition). However, this is sometimes not enough to suppress the tonal character of the lowband, giving the highband a repetitive "buzzy" sound character. Another problem occurs when two harmonic series are mixed, one with high harmonic density (low pitch) and the other with lower harmonic density (high pitch). If the high-pitched harmonic series dominates over the other in the lowband but not in the highband, the HFR causes the harmonics of the high-pitched signal to dominate the highband, causing the high band to sound "metallic" compared to the original. In high frequency reconstruction methods such as SBR, it is possible to individually control the level of the generated harmonic series, i.e. second order harmonics or third order harmonics. This is however not enough to correct the problems described above. In other implementations a constant high order of spectral whitening is introduced during the spectral envelope adjustment of the HFR signal. This solves the above-described problems, but introduces new artifacts on other signal excerpts that do not benefit from the high order spectral whitening.

25 SUMMARY OF THE INVENTION

30 The present invention relates to the problem of "buzziness" and "metallic"-sound that is sometimes introduced in HFR-methods. It uses a detection algorithm on the encoder side in order to estimate the preferable amount of spectral whitening to be applied in the decoder. The spectral whitening varies over time as well as over frequency, ensuring the best possibilities to control the harmonic contents of the replicated high band. The present invention can be carried out in a time-domain implementation as well as in a subband filterbank implementation. It uses linear prediction with variable predictor order as well as variable bandwidth expansion factor.

The present invention comprises the following features:

- In the encoder, estimation of the preferred degree of spectral whitening to use in the decoder, at a given time and in a given frequency band,
- 5 - In the decoder, perform spectral whitening in either the time domain or in a subband filterbank, in accordance with the whitening information transmitted from the encoder.
- The whitening in the decoder is obtained using linear prediction.
- The degree of whitening is controlled by varying the predictor order, or by varying the bandwidth expansion factor.
- 10 - In a subband filterbank: low-order predictors, effective implementation, especially in a system where a filterbank already is used for envelope adjustment.

BRIEF DESCRIPTION OF THE DRAWINGS

- 15 The present invention will now be described by way of illustrative examples, not limiting the scope or spirit of the invention, with reference to the accompanying drawings, in which:

Fig. 1 illustrates bandwidth expansion of an LPC spectrum.

Fig. 2a illustrates a worst case signal according to the present invention.

Fig. 2b illustrates the autocorrelation for the highband and lowband of the worst case signal.

- 20 Fig. 3 illustrates a time domain implementation of the adaptive whitening in the decoder.

Fig. 4 illustrates a subband filterbank implementation of the adaptive whitening in the decoder.

Fig. 5 illustrates an encoder implementation of the present invention.

Fig. 6 illustrates a decoder implementation of the present invention.

25

DESCRIPTION OF PREFERRED EMBODIMENTS

The below-described embodiments are merely illustrative for the principles of the present invention for improvement of high frequency reconstruction systems. It is understood that modifications and variations of the arrangements and the details described herein will be apparent to others skilled in the art. It is the intent, therefore, to be limited only by the scope of the impending patent claims and not by the specific details presented by way of description and explanation of the embodiments herein.

Linear prediction based spectral whitening

When adjusting a spectral envelope of a signal to a given spectral envelope a certain amount of whitening is always applied. This since if the desired spectral envelope is described by $H_{envRef}(f)$ and the spectral envelope of the current signal segment $H_{envCur}(f)$ the filter function applied is

35

$$H(f) = \frac{H_{\text{envRef}}(f)}{H_{\text{envCur}}(f)} \quad \text{Eq. 1.}$$

It is however not necessary to have the same frequency resolution for $H_{\text{envRef}}(f)$ as for $H_{\text{envCur}}(f)$. The current invention uses adaptive frequency resolution of $H_{\text{envCur}}(f)$ for envelope adjustment of HFR signals. The variable frequency resolution is obtained by using LPC of different order and with different bandwidth expansion factors for different frequency bands, in the time domain or in the filterbank domain, as will be explained and outlined below.

Assume the input signal segment can be represented by a time varying digital filter whose steady-state system function is of the form

$$H(z) = \frac{S(z)}{U(z)} = \frac{G}{1 - \sum_{k=1}^p a_k z^{-k}} \quad \text{Eq. 2,}$$

where $U(z)$ is the excitation signal (assumed to be a pulse train or noise) and $S(z)$ is the current input signal segment. The input signal $s(n)$ to the HFR generator in the decoder can be predicted according to

$$\tilde{s}(n) = \sum_{k=1}^p \alpha_k s(n-k) \quad \text{Eq. 3.}$$

The prediction error can be defined as

$$e(n) = s(n) - \tilde{s}(n) = s(n) - \sum_{k=1}^p \alpha_k s(n-k) \quad \text{Eq. 4.}$$

Hence the prediction error sequence is the output of a system whose transfer function is

$$A(z) = 1 - \sum_{k=1}^p \alpha_k z^{-k} \quad \text{Eq. 5.}$$

This gives an estimate of the system transfer function

$$H(z) = \frac{G}{A(z)} \quad \text{Eq. 6.}$$

In order to spectrally whiten the signal segment it can be filtered with the inverse of $H(z)$. The degree of whitening can be controlled by varying the predictor order, limiting the order of the polynomial $A(z)$, thus limiting the amount of fine structure that can be described by $H(z)$, or by applying a bandwidth expansion factor to the polynomial $A(z)$. If the bandwidth expansion factor is ρ , the polynomial $A(z)$ evaluates to

$$A(\rho z) = a_0 z^0 \rho^0 + a_1 z^1 \rho^1 + a_2 z^2 \rho^2 + \dots + a_p z^p \rho^p \quad \text{Eq. 7.}$$

This expands the bandwidth of the formants estimated by $H(z)$ according to Fig. 1. Thus the inverse filter at a given time can be described as

$$H_{inv}(z, p, \rho) = \frac{1 - \sum_{k=1}^p \alpha_k (z\rho)^{-k}}{G} \quad \text{Eq.8.}$$

The coefficients α_k can be obtained in different manners, e.g. the autocorrelation method or the covariance method [Digital Signal Processing, principles, algorithm and applications, Proakis & Manolakis, Prentice Hall]. The gain factor G can be set to one if the H_{inv} is used as a pre-whitening filter followed by a regular envelope adjustment. It is common practice to add some sort of relaxation to the estimate in order to ensure stability of the system. When using the autocorrelation method this is easily accomplished by offsetting the zero-lag value of the correlation vector. This is equivalent to addition of white noise to the signal used to estimate $A(z)$. The parameters p and ρ are calculated based on information transmitted from the encoder.

An alternative to bandwidth expansion is described by:

$$A_b(Z) = 1 - b + b \cdot A(z) \quad \text{Eq. 9,}$$

where b is the whitening factor. This yields the whitening filter according to:

$$H_{inv}(z, p, b) = \frac{1 - b + b \cdot (1 - \sum_{k=1}^p \alpha_k (z)^{-k})}{G} \quad \text{Eq. 10.}$$

Here it is evident that for $b = 1$ Eq. 10 evaluates to Eq. 8 with $\rho = 1$, and for $b = 0$ Eq. 10 evaluates to a constant non frequency selective gain factor.

The detector on the encoder side

In the encoder a detector is used to assess the best degree of whitening (LPC order, bandwidth expansion factor and whitening factor) to be used in the decoder in order to obtain a highband as similar to the original given the currently used HFR method. Several approaches can be used in order to obtain a proper estimate of the degree of whitening to be used in the decoder. These include e.g. analysis by synthesis. The detector can be made to estimate the most suitable order of the LPC used in the decoder and different bandwidth expansion factors for different frequency bands.

One approach uses auto correlation to estimate the appropriate amount of whitening. The detector estimates the autocorrelation functions for the source range and the target range, Fig 2a. Here a worst case signal is described, a harmonic series in the low band and white noise in the highband. The different autocorrelation functions are displayed in Fig 2b. Here it is evident that the lowband is highly correlated whilst the highband is not. The

maximum correlation for all lags larger than a minimum lag, are obtained for both the highband and the lowband. The quotient of the two is used to calculate the optimal bandwidth expansion factor. In the current implementation of the above outlined system, it is preferable to use FFT:s for the computation of the correlation. The autocorrelation of a sequence $x(n)$ is defined by:

$$r_{xx}(m) = FFT^{-1}(|X(k)|^2) \quad \text{Eq. 11,}$$

where

$$X(k) = FFT(x(n)) \quad \text{Eq. 12.}$$

Since it is the objective to compare the difference of the auto correlation in the highband and the lowband the filtering can be done in the frequency domain. This yields:

$$\begin{cases} X_{Lp}(k) = X(k) \cdot H_{Lp}(k) \\ X_{Hp}(k) = X(k) \cdot H_{Hp}(k) \end{cases} \quad \text{Eq. 13,}$$

where $H_{Lp}(k)$ and $H_{Hp}(k)$ are the Fourier transforms of the LP and HP filters impulse responses. From the above the auto correlation functions for the lowband and highband can be calculated according to:

$$\begin{cases} r_{xxLp}(m) = FFT^{-1}(|X_{Lp}(k)|^2) \\ r_{xxHp}(m) = FFT^{-1}(|X_{Hp}(k)|^2) \end{cases} \quad \text{Eq. 14.}$$

The maximum value, for a lag larger than a minimum lag, for each autocorrelation vector is calculated:

$$\begin{cases} r_{MaxLp} = \max(r_{xxLp}) \quad \forall m > \text{minLag} \\ r_{MaxHp} = \max(r_{xxHp}) \quad \forall m > \text{minLag} \end{cases} \quad \text{Eq. 15.}$$

The quota can be used to for instance linearly map to a suitable bandwidth expansion factor.

A different detector approach is to obtain an estimate of the harmonic to non-harmonic signal ratio H in each subband of a complex filter bank by using very low order linear prediction for each block of subband samples. The energy of the predicted block of subband samples divided by the total energy of the block gives an estimate of H as a function of both time and frequency. The difference between highband and lowband values of H is then used to adjust the degree of whitening such that the harmonic to signal ratio of the synthetically generated highband approaches that of the original highband. Here it is advantageous to control the degree of whitening utilising the whitening factor b (Eq. 9).

Adaptive LPC-based whitening in the time domain

When performing whitening in the time domain the auto correlation method is preferred. The auto correlation method requires windowing of the input segment used to estimate the coefficients α_k , this is not the case for the covariance method. The filter used for the spectral whitening is

$$H_{inv}(z, p, \rho) = 1 - \sum_{k=1}^p \alpha_k (z\rho)^{-k} \quad \text{Eq. 16,}$$

here the gain factor G is set to one. It's beneficial to whiten the signal in the low band prior to the HFR generator, since the whitening thus can operate on a lower sampling rate. The lowband signal is windowed and whitened on a suitable time base with the predictor order and bandwidth expansion factors given by the encoder, according to Fig. 3. In the current implementation of the present invention the signal is low pass filtered 301 and decimated 302. 303 illustrate the spectral whitening. A window 306 is used to select the proper time segment for estimation of the $A(z)$ polynomial 307, 50% overlap is used. The LPC-routine extracts $A(z)$ given the currently preferred LPC-order and bandwidth expansion factor, with a suitable relaxation. A FIR filter 308 is used to inverse filter the signal segment. The whitened signal segments are upsampled 304, 305 and windowed together forming the input signal to the HFR unit.

Adaptive LPC-based whitening in a subband filter bank

The whitening can be performed effectively and robustly by using a complex filter bank. The linear prediction and the inverse filtering are then done independently for each of the subband signals produced by the filter bank. The output from a complex filter bank used for the analysis of a real valued signal can be interpreted as a set of subband signals that are oversampled a factor of two in the frequency domain. This feature is exploited to heavily reduce artifacts due to aliasing emerging from independent modifications of the subband signals, which for example inverse filtering results in.

The whitening of the subband signals is obtained through linear prediction analogous to the time domain method described above. The subband signals are however complex valued, so complex filter coefficients are used for the linear prediction as well as for the inverse filtering. The order of the linear prediction can be kept very low since the expected number of harmonic components in each frequency band is very small. In order to correspond to the same time base as the time domain LPC, the number of subband samples in each block is smaller by a factor equal to the downsampling of the filter bank. Given the low filter order and small block sizes the prediction filter coefficients are obtained from the covariance method. Although no windowing of subband signal blocks is performed prior to the computation of covariance estimates, the inner products are weighted by a positive valued window $w(n)$. The matrix elements involved for a block size of N samples for the subband signal corresponding to channel l are thus:

$$\phi_l(i, k) = \sum_{n=0}^{N-1} x_l(n-i) x_l^*(n-k) w(n), \quad 0 \leq i \leq p, \quad 1 \leq k \leq p. \quad \text{Eq. 17}$$

Here, $x_l(n)$ is the subband signal, $x_l^*(n)$ is its complex conjugate, and p is the order of linear prediction. The purpose of this internal windowing is mainly to enhance the whitening of tonal signals with time varying frequency content. The coefficients obtained from these covariances are used to filter the subband signal block $[x_l(0), x_l(1), \dots, x_l(N-1)]$. Filter coefficient calculation and whitening can be performed on a block by block basis

using subband sample time step L , which is smaller than the block length N . The whitened blocks should be added together using appropriate synthesis windowing.

Feeding a maximally decimated filterbank with an input signal consisting of white gaussian noise will produce subband signals with white spectral density. In a complex filter bank, the subband signals are oversampled by a factor of two as mentioned above. Feeding the oversampled filterbank with white noise gives subband signals with coloured spectral density. This is due to the effects of the frequency responses of the analysis filters. The LPC predictors in the filterbank channels will track the filter characteristics in the case of noise-like input signals. This is an unwanted feature, and benefits from compensation. A possible solution is pre-filtering of the input signals to the linear predictors. The pre-filtering should be an inverse, or an approximation of the inverse, of the analysis filters, in order to compensate for the frequency responses of the analysis filters. The whitening filters are fed with the original subband signals, as described above. Fig. 4 illustrates the whitening process of a subband signal. The complex-valued signal corresponding to channel l is fed to the pre-filtering block 401, and subsequently to a delay chain which depth depends on the filter order 402. The delayed signals and their conjugates 403 are fed to the linear prediction block 404, where the complex-valued coefficients are calculated. The coefficients from every L :th calculation are kept by the decimator 405. The subband signal are finally filtered through the filterblock 406, where the predicted coefficients are used and updated for every L :th samples.

Practical implementations

The present invention can be implemented in both hardware chips and DSPs, for various kinds of systems, for storage or transmission of signals, analogue or digital, using arbitrary codecs. Fig. 5 and Fig. 6 shows a possible implementation of the present invention. In Fig. 5 the encoder side is displayed. The analogue input signal is fed to the A/D converter 501, and to an arbitrary audio coder, 502, as well as the inverse filtering level estimation unit 503, and an envelope extraction unit 504. The coded information is multiplexed into a serial bitstream, 505, and transmitted or stored. In Fig. 6 a typical decoder implementation is displayed. The serial bitstream is de-multiplexed, 601, and the envelope data is decoded, 602, i.e. the spectral envelope of the high-band. The de-multiplexed source coded signal is decoded using an arbitrary audio decoder, 603. The decoded signal is fed to the spectral whitening unit 604, which performs the adaptive spectral whitening. The spectrally whitened signal is fed to an arbitrary HFR unit, 605, and to the envelope adjuster 606. The output from the envelope adjuster is combined with the decoded signal fed through a delay, 607. Finally, the digital output is converted back to an analogue waveform 608.

CLAIMS

1. A method for enhancement of source coding systems using high-frequency reconstruction, where said source coding system comprises an encoder representing all operations performed prior to storage or transmission, and a decoder representing all operations performed after storage or transmission, characterised by:

at said encoder, estimating the required amount of spectral whitening at a given time and frequency, and transmitting information on said amount of spectral whitening from said encoder to said decoder;

at said decoder, adaptively, spectrally whiten a signal prior to HFR or after HFR, according to the whitening information obtained from said encoder.

2. A method according to claim 1, characterised in that said spectral whitening is performed in the time domain.

3. A method according to claim 1, characterised in that said spectral whitening is performed in a subband filterbank.

4. A method according to claim 1, characterised in that said estimation of required amount of spectral whitening is done using analysis by synthesis.

5. A method according to claim 1, characterised in that said estimation of required amount of spectral whitening is done by comparative study of the autocorrelation in the lowband and the highband.

6. A method according to claim 1, characterised in that the amount of spectral whitening is controlled by the LPC predictor order.

7. A method according to claim 1, characterised in that the amount of spectral whitening is controlled by the bandwidth expansion factor of the LPC polynomial.

8. A method according to claim 1, characterised in that the amount of spectral whitening is controlled by the whitening factor b .

9. A method according to claim 3, characterised in that different bandwidth expansion factors are used for different filterbank bands.

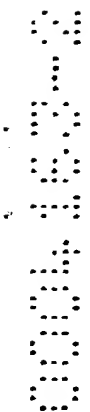
10. A method according to claim 3, characterised in that pre-filtering is included in the LPC estimation in order to compensate for the characteristic of the filterbank analysis filters.

means for adaptively, spectrally whiten a signal prior to HFR or after HFR, in said decoder according to the

5 whitening information obtained from said encoder.

ABSTRACT

The present invention proposes new methods and an apparatus for enhancement of source coding systems utilising high frequency reconstruction (HFR). It uses a detection algorithm on the encoder side in order to estimate the correct amount of spectral whitening to be applied in the decoder. The spectral whitening varies over time as well as over frequency, ensuring the best possibilities to control the harmonic contents of the replicated high band. The present invention is applicable to both speech coding and natural audio coding systems.



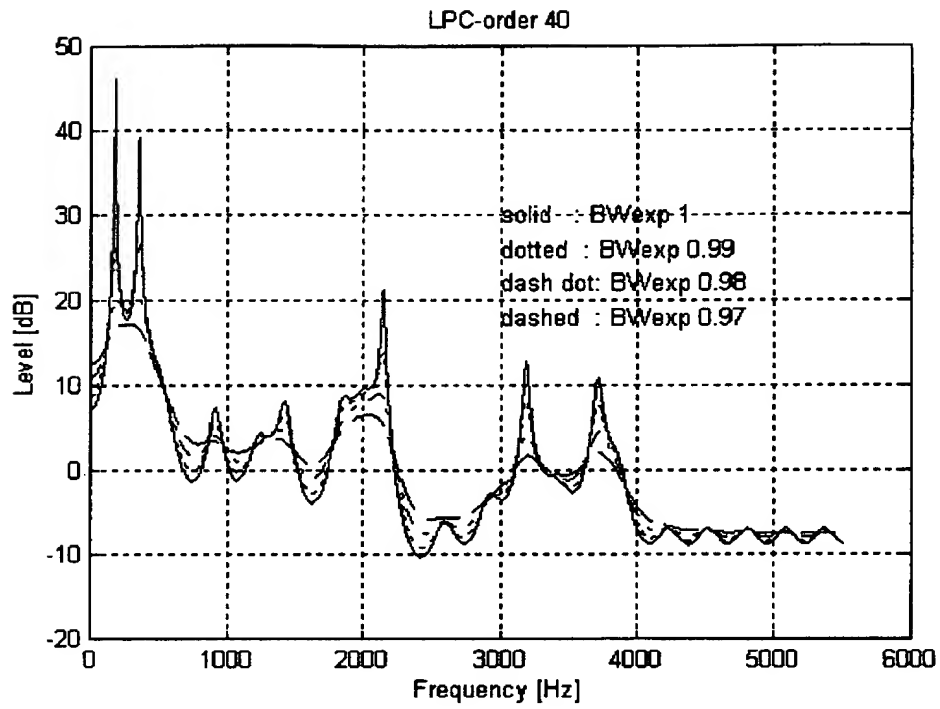


Fig. 1

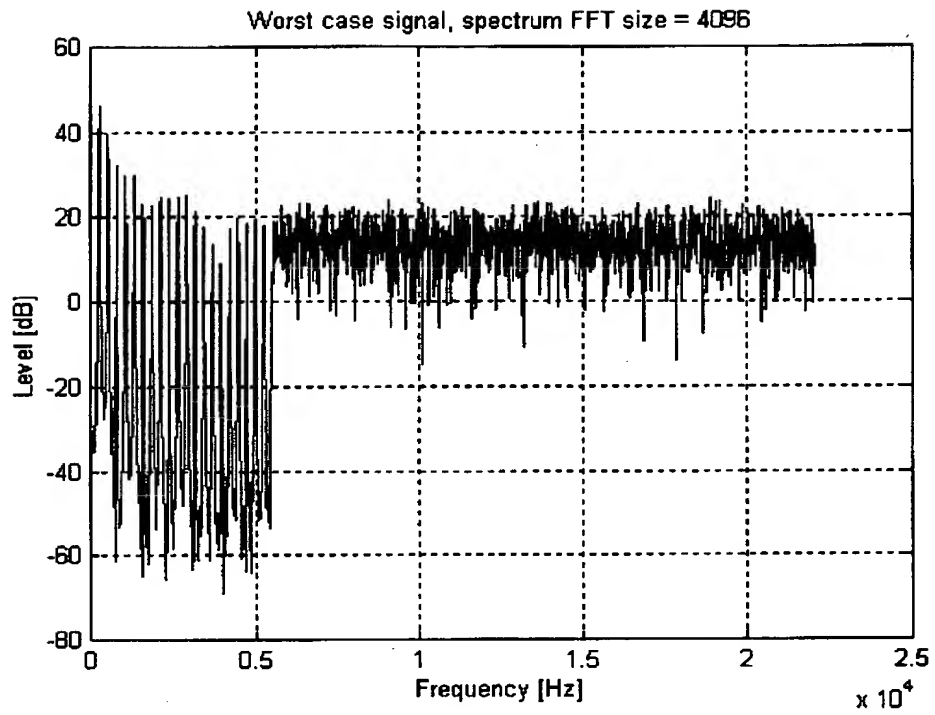


Fig. 2a

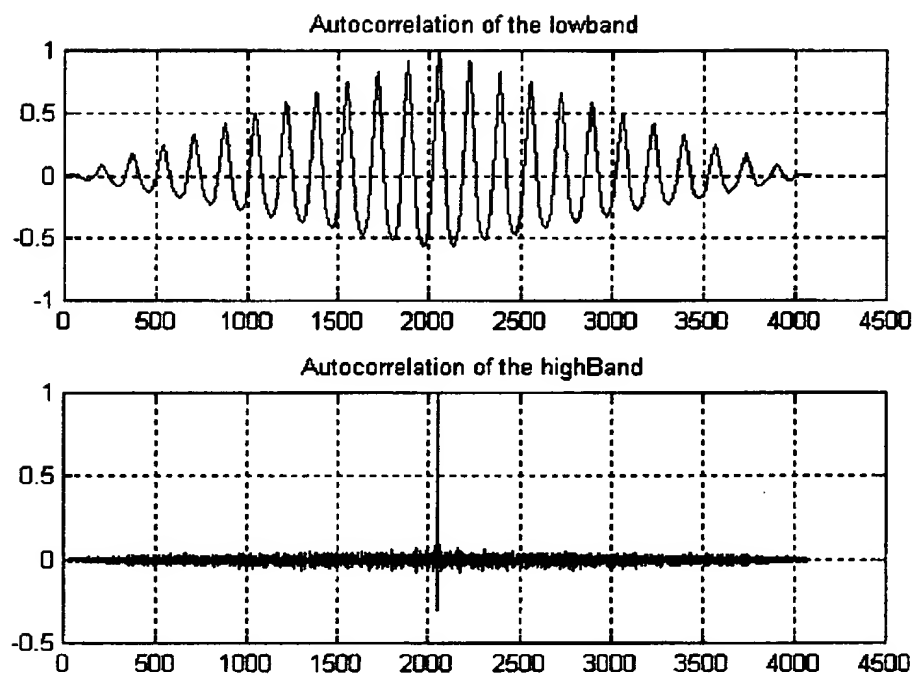


Fig. 2b

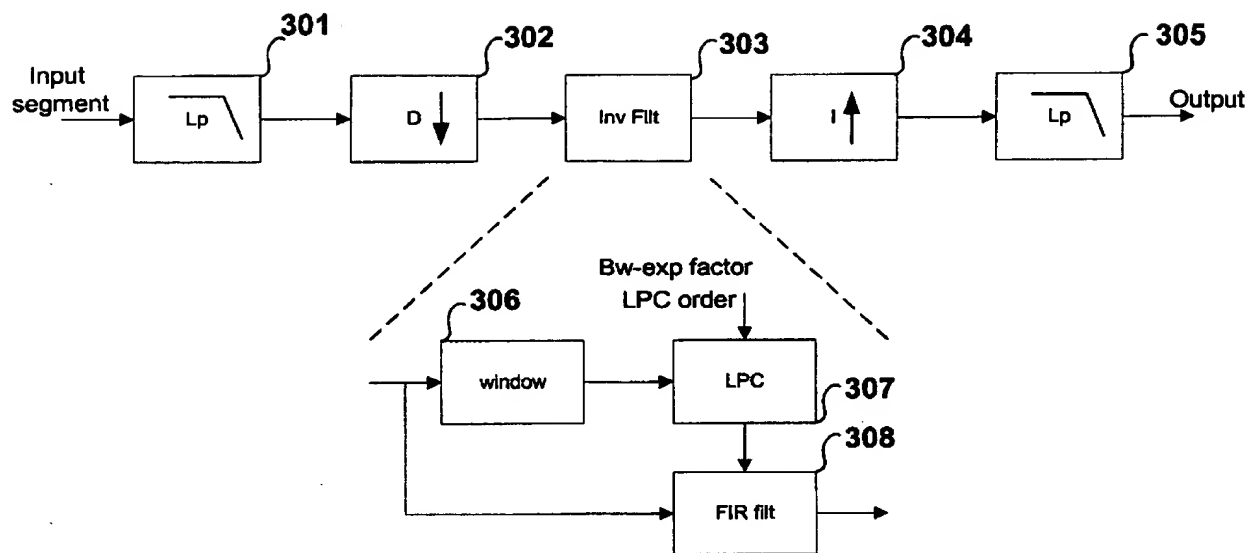


Fig. 3

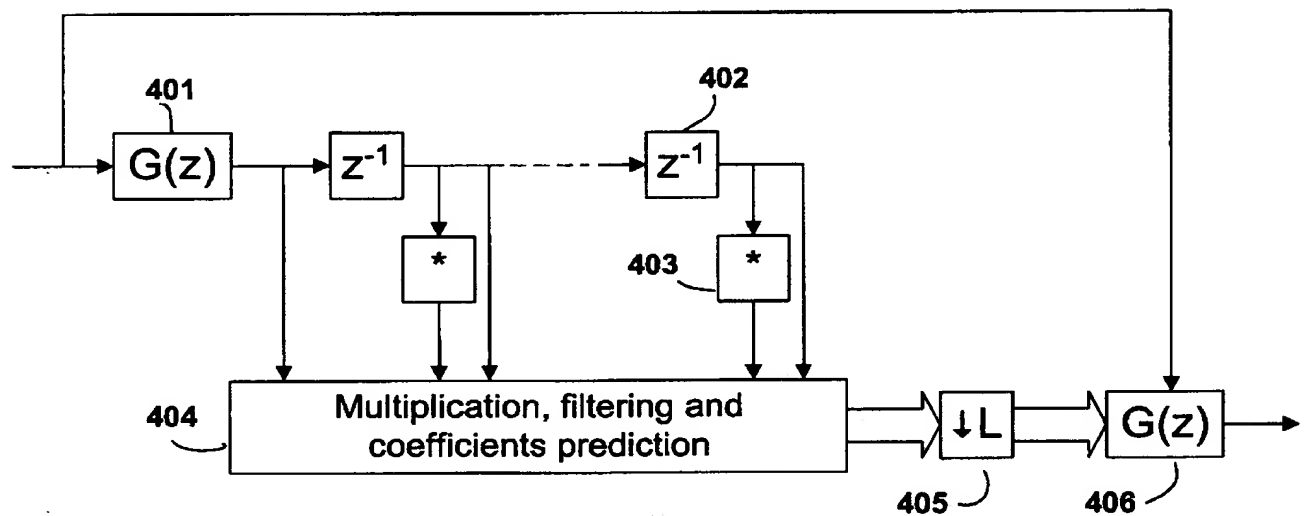


Fig. 4

0004153-3

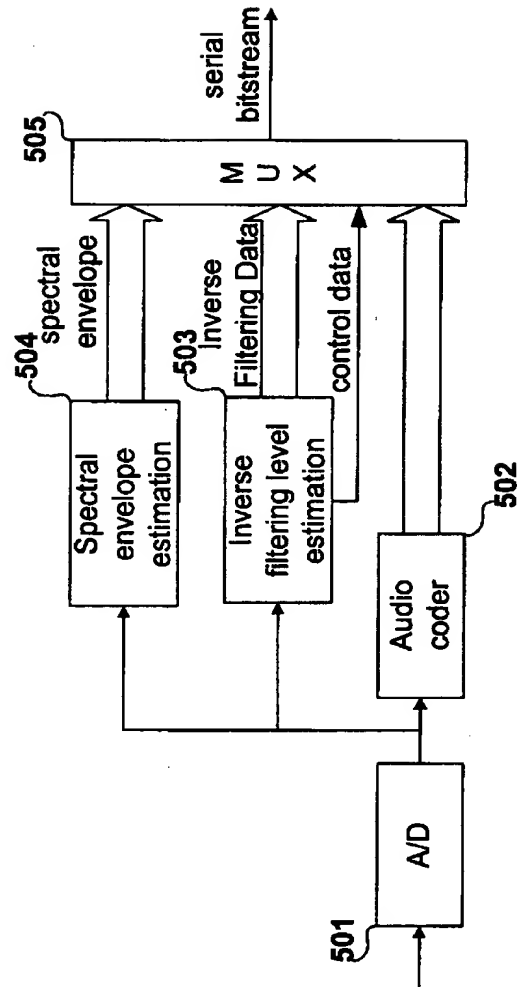


Fig. 5

00041533-2

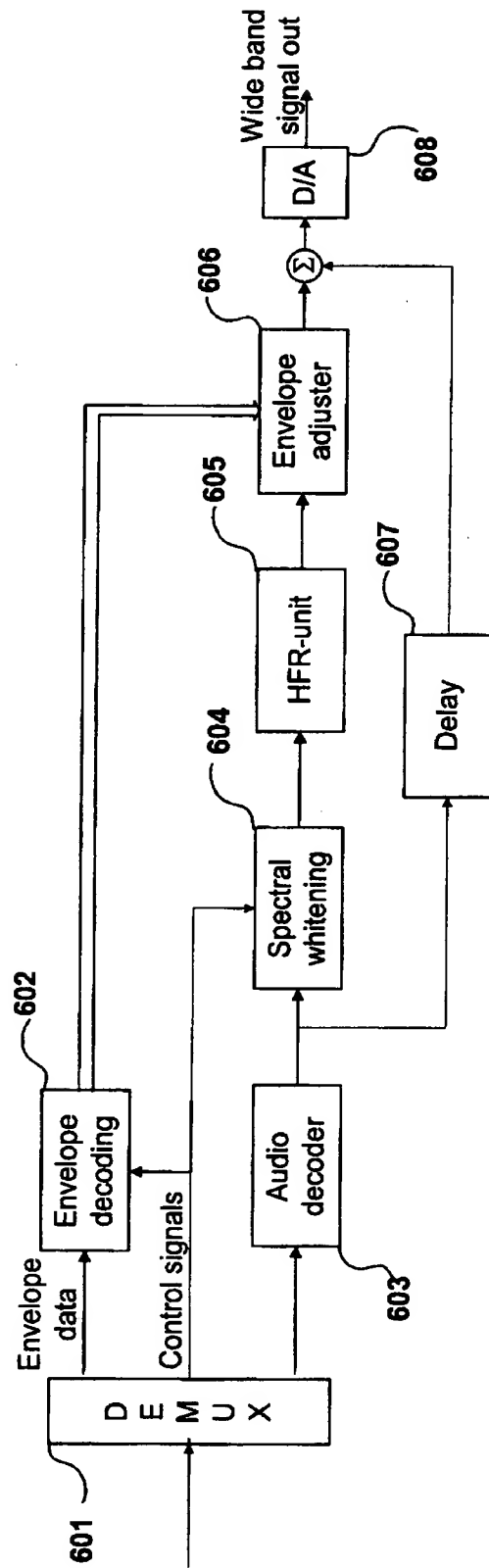


Fig. 6

Under the Paperwork Reduction Act of 1995, no person are required to respond to a collection of information unless it displays a valid OMB control number.

**PROCESSING FEE
Under 37 CFR 1.17(i)
TRANSMITTAL**

(Fees are subject to annual revision)

Send completed form to:Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Application Number	09/987,475-Conf. #007069
Filing Date	November 14, 2001
First Named Inventor	Kristofer KJORLING
Art Unit	2654
Examiner Name	V. B. Chawan
Attorney Docket Number	4441-0109P

Enclosed is a paper filed under 37 CFR 1.55 that requires a processing fee (37 CFR 1.17(i)).
Payment of \$ 130.00 is enclosed.

This form should be included with the above-mentioned paper and faxed or mailed to the Office using the appropriate Mail Stop, if applicable. For transmittal of petition fees under 37 CFR 1.17(f), (g) or (h), see form PTO/SB/17p.

Payment of Fees (small entity amounts are NOT available for the processing fees)

☒ The Commissioner is hereby authorized to charge the following fees to Deposit Account No. 02-2448 :
☐ processing fee under 37 CFR 1.17(i) ☒ any deficiency of fees and credit of any overpayments

Enclose a duplicative copy of this form for fee processing.

☒ Check in the amount of \$ 130.00 is enclosed.

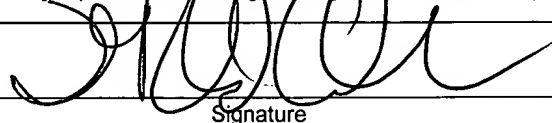
☐ Payment by credit card (Form PTO-2038 or equivalent enclosed). Do not provide credit card information on this form.

Processing Fees under 37 CFR 1.17(i): Fee \$130

**Fee Code 1808 for all,
Except for \$1.221 papers (Fee Code 1803)**

For papers filed under:

- § 1.28(c)(3) – for processing a non-itemized fee deficiency based on an error in small entity status.
- § 1.41 – for supplying the name or names of the inventor or inventors after the filing date without an oath or declaration as prescribed by § 1.63, except in provisional applications.
- § 1.48 – for correcting inventorship, except in provisional applications.
- § 1.52(d) – for processing a nonprovisional application filed with a specification in a language other than English.
- § 1.53(b)(3) – to convert a provisional application filed under § 1.53(c) into a nonprovisional application under § 1.53(b).
- § 1.55 – for entry of late priority papers.
- § 1.71(g)(2) – to enter an amendment to the specification for purposes of 35 U.S.C. 103(c)(2) if not filed within the cited time periods
- § 1.99(e) – for processing a belated submission under § 1.99.
- § 1.103(b) – for requesting limited suspension of action, continued prosecution application (§ 1.53(d)).
- § 1.103(c) – for requesting limited suspension of action, request for continued examination (§ 1.114).
- § 1.103(d) – for requesting deferred examination of an application.
- § 1.217 – for processing a redacted copy of a paper submitted in the file of an application in which a redacted copy was submitted for the patent application publication.
- § 1.221 – for requesting voluntary publication or republication of an application. **Fee Code 1803**
- § 1.291(c)(5) – for processing a second or subsequent protest by the same real party in interest.
- § 1.497(d) – for filing an oath or declaration pursuant to 35 U.S.C. 371(c)(4) naming an inventive entity different from the inventive entity set forth in the international stage.
- § 3.81 – for a patent to issue to assignee, assignment submitted after payment of the issue fee.


Signature

November 9, 2005
Date

D. Richard Anderson
Typed or printed name

40,439

1 Registration No. 09987475

01 FC:1464

130 00 0P